

Anthropogenic heat flux: advisable spatial resolutions when input data are scarce

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Gabey, A. M., Grimmond, C. S. B. and Capel-Timms, I. (2019) Anthropogenic heat flux: advisable spatial resolutions when input data are scarce. *Theoretical and Applied Climatology*, 135 (1-2). pp. 791-807. ISSN 1434-4483 doi: <https://doi.org/10.1007/s00704-018-2367-y> Available at <https://centaur.reading.ac.uk/75551/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1007/s00704-018-2367-y>

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Anthropogenic heat flux: advisable spatial resolutions when input data are scarce

A. M. Gabey¹ · C. S. B. Grimmond¹ · I. Capel-Timms¹

Received: 3 August 2017 / Accepted: 2 January 2018
© The Author(s) 2018. This article is an open access publication

Abstract

Anthropogenic heat flux (Q_F) may be significant in cities, especially under low solar irradiance and at night. It is of interest to many practitioners including meteorologists, city planners and climatologists. Q_F estimates at fine temporal and spatial resolution can be derived from models that use varying amounts of empirical data. This study compares simple and detailed models in a European megacity (London) at 500 m spatial resolution. The simple model (LQF) uses spatially resolved population data and national energy statistics. The detailed model (GQF) additionally uses local energy, road network and workday population data. The Fractions Skill Score (FSS) and bias are used to rate the skill with which the simple model reproduces the spatial patterns and magnitudes of Q_F , and its sub-components, from the detailed model. LQF skill was consistently good across 90% of the city, away from the centre and major roads. The remaining 10% contained elevated emissions and “hot spots” representing 30–40% of the total city-wide energy. This structure was lost because it requires workday population, spatially resolved building energy consumption and/or road network data. Daily total building and traffic energy consumption estimates from national data were within $\pm 40\%$ of local values. Progressively coarser spatial resolutions to 5 km improved skill for total Q_F , but important features (hot spots, transport network) were lost at all resolutions when residential population controlled spatial variations. The results demonstrate that simple Q_F models should be applied with conservative spatial resolution in cities that, like London, exhibit time-varying energy use patterns.

1 Introduction

Anthropogenic heat flux (Q_F) can be a substantial input into the urban energy balance, especially in mid- and high-latitude cities during the autumn, winter and early spring when it matches or even exceeds the incoming radiant flux in some areas of a city (e.g. Kikegawa et al. 2003; Hamilton et al. 2009). Despite its magnitude, Q_F is difficult to measure directly because it is incorporated into measured sensible, latent, storage and long-wave radiative fluxes. Q_F is therefore commonly estimated using empirical models (e.g. Sailor and Lu 2004; Smith et al. 2009; Allen et al. 2011; Iamarino et al.

2012) that simulate the metabolic ($Q_{F,M}$), transportation ($Q_{F,T}$) and building ($Q_{F,B}$) components of Q_F .

Modelling methodologies for Q_F are either top-down, bottom-up or a combination of the two. Top-down models (e.g. Sailor and Lu 2004; Allen et al. 2011) combine energy consumption or statistical parameters summarising large areas with supplementary data to produce estimates at finer spatial and temporal scales. Bottom-up approaches model emissions from individual buildings (e.g. Bueno et al. 2012) or road links (Smith et al. 2009) and apply these across large areas. A combination of the two (e.g. Iamarino et al. 2012) may be used to balance the desirable accuracy of bottom-up methods with the practicality of the top-down approach, given comprehensive input data are not always available.

Top-down methods are appealing as they (i) allow studies in regions where data are scarce and (ii) are applicable to large geographical areas that can be incorporated into numerical weather prediction or global/regional climate models (e.g. Flanner 2009), and (iii) their relative simplicity makes them computationally tractable. These methods generally use a spatial dataset of local population density (e.g. Sailor and Lu 2004; Allen et al. 2011) to distribute national/regional per-

✉ A. M. Gabey
a.m.gabey@reading.ac.uk

¹ Department of Meteorology, University of Reading, Reading, UK

capita energy consumption and vehicle ownership statistics within the domain of interest. The availability of world-wide gridded population datasets (e.g. GPWv4—Center for International Earth Science Information Network 2016; GHS—Pesaresi 2015) means theoretically Q_F can be modelled at sub-kilometre resolution. This makes its use in urban energy balance models appealing, but the potential for inaccuracy rises with finer resolution since the input data may be insufficient to reflect energy consumption across a city at such scales.

A Q_F model is considered skilful in this analysis if the resulting spatial distributions accurately reproduce a reference model or dataset in different Q_F regimes: areas with (i) zero emission, (ii) “hot spots” (regions with the most intense emission) and (iii) intermediate intensities. Different Q_F regimes and/or components are of interest for different applications, so a skill score must be able to objectively compare them by incorporating the abundance and spatial and temporal accuracy of emissions in each regime.

Previous studies comparing Q_F model results use subjective assessments of spatial assumptions. For example, Allen et al. (2011) note that using residential population datasets to attribute urban emissions may not accurately represent the true energy use pattern. Quantitative comparisons focus on extreme values or large-area averages rather than spatial accuracy or bias: Sailor et al. (2015) compare maximum values and city-wide averages; Dong et al. (2017) discuss global mean values (and note the limitation of this approach); Lindberg et al. (2013) demonstrate how the origin and resolution of residential population data affect cumulative Q_F . While broadly informative, these strategies leave an incomplete understanding of the merits and limitations of different modelling approaches.

In this study, the Fractions Skill Score (FSS) and supplementary analyses are used to evaluate skill objectively and to establish the spatial resolution and time(s) of day at which a simple model produces skilful estimates. A detailed Q_F model (GreaterQF; Iamarino et al. 2012) is used as a reference standard against which the skill of a top-down model (LUCY; Allen et al. 2011; Lindberg et al. 2013) is rated.

The LUCY and GreaterQF models have been re-implemented (under the Urban Multiscale Environmental Predictor (UMEP) suite—Lindberg et al. 2018) to interface with the Quantum GIS Geographical Information System (2016) to use standard GIS datasets as inputs. The models are termed “LQF” (LUCY) and “GQF” (GreaterQF) hereafter to emphasise the new implementations presented here, which make it easier to update input data to represent different historical or hypothetical conditions.

Information about the FSS and models is provided in Sect. 2, and the skill with which LQF predicts Q_F , $Q_{F,T}$, $Q_{F,M}$ and $Q_{F,B}$ is discussed in Sect. 3, along with further analyses that apply different spatial considerations.

2 Methods

2.1 Overview of models

The modifications and improvements in LQF and GQF include pre-processing of input data to attach weightings to each grid cell in the model domain, with grid cell size defined by the spatial resolution of population count data used as input. The weightings are used to apportion vehicle activity, building energy consumption and metabolic activity, from which Q_F emissions are estimated (Appendix 2). Weightings in LQF are based only on residential population density, while in GQF the weightings combine population density, energy consumption (broken down by fuel and consumer type), total road length (by road class) and traffic flow (by road link and vehicle type).

LQF contains a user-editable georeferenced database of historic national energy consumption, vehicle ownership (car, motorcycle and freight/bus) and total population statistics. Each grid cell receives a share of the energy consumption and vehicles based on the proportion of national population it contains. Building energy use depends on mean daily air temperature recorded for the study area via an empirical scaling function (Lindberg et al. 2013), which does not conserve the annual per-capita energy consumption derived from the database. Metabolism is simulated using the residential population at all times of day. Hourly profiles of building energy use, metabolic activity and road traffic variations are applied to capture the diurnal variations of these energy sources.

All of the energy input into LQF and GQF is assumed to be released as latent, sensible and/or wastewater Q_F . GQF uses city-scale data to estimate emissions: Building emissions are calculated using spatially resolved annual energy consumption (electricity and gas), which distinguish non-domestic and domestic consumption within the city. These are disaggregated to a common spatial resolution using workday and residential population data, respectively. Daily building energy consumption variations are generated based on electricity and gas demand data, and separate non-domestic and domestic half-hourly profiles are applied to capture diurnal variations. Unlike in LQF, annual per-capita building energy consumption is conserved in GQF. Road traffic fuel use for eight vehicle types is built up from a detailed map of road links, and vehicle-specific profiles are applied to elicit diurnal variations. Metabolic emissions are represented using the residential population at night and the workday population on working days, with a transitional period between them. Each grid cell therefore contains different proportions of transport, building and metabolic Q_F .

2.1.1 Model configuration

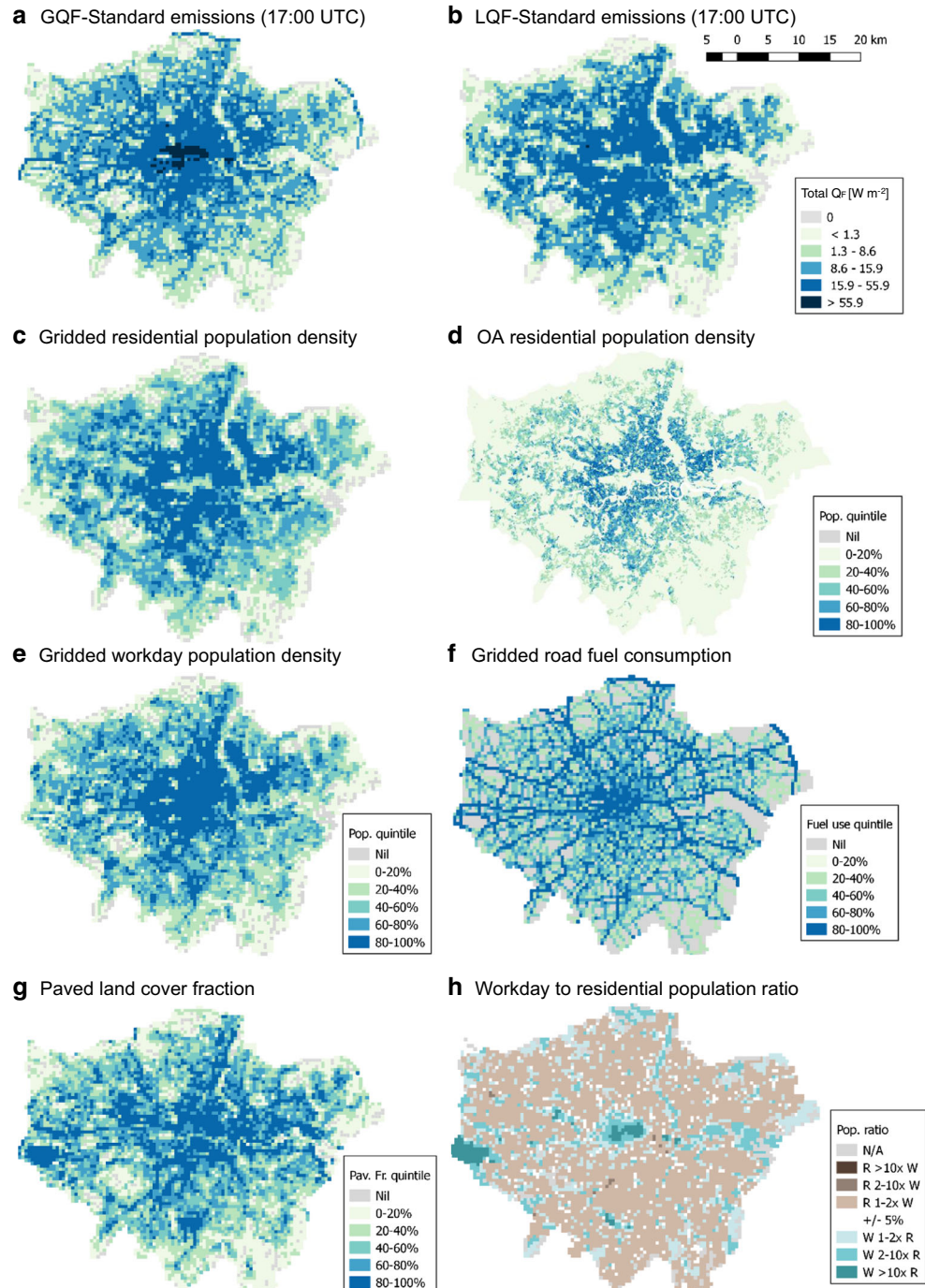
LQF and GQF models were run for Greater London, UK. The study area (residential population 8.2 M, workday population 8.67 M—Office for National Statistics 2016) extends $60 \times$

45 km at its widest points and contains diverse land uses and substantial population variations. Data requirements are met for both of the models. A substantial change in population distribution occurs between night (Fig. 1c) and working days (Fig. 1e). This leads to large workday enhancements in the city centre and far west and modest but widespread population decreases in the majority of the city (Fig. 1h). The road network contains arterial routes, an inner ring-road in the northern half of the city and an outer highway partly intersecting the edge of the study area. These features are visible in the daily

mean vehicle fuel consumption estimate produced by the GQF pre-processing steps (Fig. 1f).

The spatial accuracy of LQF is evaluated on a working day: Tuesday, 5 May 2015. This falls during British Summer Time (UTC + 1, which is explicitly captured in both models). On this date, the mean air temperature recorded in central London was 9.03 °C, cool enough to trigger building heating calculations in LQF. Year-long model runs were also performed to assess LQF predictions of daily building energy consumption relative to GQF.

Fig. 1 **a, b** Total Q_F from GQF and LQF at 500 m spatial resolution at 17:00 UTC on May 5, 2015 (same colour scale). **c, d** The residential population density respectively by output area (OA) and after gridding to 500 m resolution. **e** Gridded workday population. **f** Gridded total road vehicle fuel consumption estimated in GQF (Appendix 2). **g** Paved land cover fraction and **(h)** the ratio of workday (W) to residential (R) population at 500 m resolution, coloured based on which dominates. Panels (c) to (g) are shaded by their respective quintiles. Grey cells contain zero population, emissions or fuel consumption



GQF and LQF model parameters are summarised in Appendix 1, along with details of the input data sources used. LQF was assigned one transport and one building energy consumption diurnal profile, which are averages of those used in GQF, weighted respectively by the abundance of each vehicle type (outer London; DEFRA 2014) and energy consumer (city-wide values; DBEIS 2016). The profiles and GQF outputs were temporally averaged to match the 60-min time resolution of LQF. Both models were configured to generate Q_F at 500 m spatial resolution to approximately match that of the world-wide population datasets such as Gridded Population of the World v4 (GPWv4; Center for International Earth Science Information Network 2016).

Modifications are applied to LQF and GQF to quantify the change in skill from varying spatial input datasets (Table 1 contains a breakdown of input data used for each variant):

- GQF and LQF: standard model configurations
- “LQF-Paved”: as LQF, but using remote-sensed paved area fraction data to estimate the spatial distribution of $Q_{F,T}$
- “GQF-Simplified”: spatially resolved energy consumption input to GQF is replaced with city-wide values to document the effect on skill.

Hourly spatial distributions of total Q_F were produced by GQF and LQF for the whole of 5 May 2015. Results from the 16:00 to 17:00 UTC time step are shown in Fig. 1a, b as examples. Emissions range from zero to more than 100 W m^{-2} over the city, with different spatial distributions resulting from each model treatment.

2.1.2 Input data

Per-capita building energy use Per-capita building energy consumption is estimated by LQF using UK total population (62.036 M) and energy consumption statistics (Allen et al. 2011; Lindberg et al. 2013), both from 2010. In comparison, London-specific energy consumption (2014) and population (2011) data are used to calculate per-capita building energy use in GQF. Appendix 1 contains full listings of LQF and GQF input data sources and parameters.

The UK total national primary consumption (PC) in the existing LUCY database was 3.03 times greater per-capita than the city-specific value input into GQF (Fig. 2a) and 2.92 times greater than an alternative estimate of London energy consumption with transport excluded from 2014 (DECC 2014). This disparity is dominated by the use of the total primary energy supply (IEA 2016) in LUCY, which includes energy production minus exports and losses rather than consumption.

The 2010 UK Total Final Energy Consumption minus transport fuel, or NTFC (IEA 2016), differed from the GQF

value by only 1.2% and is therefore used in LQF here instead. Changes in consumption between 2010 (LQF input) and 2014 (GQF input) are also likely to contribute to the difference between model inputs, but a full evaluation of top-down energy consumption estimates is beyond the scope of this work.

Transport fuel consumption Published aggregated total fuel consumptions for the UK and London were unsuitable for modelling road transport emissions because they included aviation and rail transport. Instead, GQF-estimated per-capita road transport fuel consumption is based on a map of traffic volume per road link from 2013 (London Datastore 2014) and is 1.4 times greater than the equivalent LQF estimate (Fig. 2b) from UK per-capita vehicle ownership statistics (Worldmapper 2006).

Day-to-day variations Daily mean temperature data for LQF was measured in Central London during the year 2015 by a Davis Vantage Pro 2 Plus weather station atop Barbican Cromwell Tower, a 42-storey residential block (51.521°N , -0.0930°E , 145 m a.s.l). Daily gas and electricity demand variations for GQF (National Grid 2016a, b) were also sourced for 2015. Section 3.2.1 compares the day-to-day variation in building energy available in each model.

LQF road traffic volume is reduced by 20% during weekends and public holidays. Diurnal and day-of-week traffic variations in GQF are governed by a single week-long profile at 30 min resolution.

Population data Spatially resolved Greater London population maps from the 2011 UK census (ONS 2013) are used to allocate population and thus energy consumption across the spatial domains of both models. To represent the effect of grid resolution on skill, gridded population data are produced from the census output area (OA) spatial units that vary in area (Fig. 1d). The populations were redistributed to a regular grid based on the plan area index of buildings according to satellite-derived land cover classification data (Marconcini et al. 2017), using the process described in Appendix 2. This yields residential and workday populations downsampled to 500 m resolution (Fig. 1c, e) for use in the two models.

2.2 Fractions skill score

The total Q_F output by each model¹ at 17:00 UTC (Fig. 1a, b) demonstrates several features that the evaluation must capture:

1. The different frequency and spatial arrangement of hot spots (emissions typically ranging from ~ 50 to over

¹ See http://urban-climate.net/umep/UMEP_Manual/_Tutorials/LQF for a tutorial that demonstrates how to use the models to replicate these outputs

Table 1 Summary of input data attributes required by the different GQF and LQF configurations, with decreasingly detailed configurations further to the right. More entries in a given column imply a more detailed configuration. See Appendix 1 for details regarding the datasets used

More detailed		Less detailed			
Input data	Components	GQF-standard	GQF-simplified	LQF Paved	LQF
Annual building energy consumption	Residential electricity	Spatially resolved	City-wide value	Total energy consumption (from national data)	Total energy consumption (from national data)
Residential gas		Spatially resolved		City-wide value	
Industrial gas		Spatially resolved		City-wide value	
Industrial electricity		Spatially resolved		City-wide value	
Energy consumption diurnal profiles	Residential electricity	Dedicated profile	Dedicated profile	One profile for all components	One profile for all components
Industrial electricity		Dedicated profile		Dedicated profile	
Residential gas		Dedicated profile		Dedicated profile	
Industrial gas		Dedicated profile		Dedicated profile	
Relative daily energy demand	Gas	Measured demand	Measured demand	Total demand estimated using temperature	Total demand estimated using temperature
Electricity		Measured demand		Measured demand	
Spatial population density (times used)	Night, weekend	Residential	Residential	Residential	Residential
Working day	Workday	Workday	Residential	Residential	
Road network layout	–	Road link map, multiple road classes	Paved land cover fraction	Population density	
Traffic heat estimate	–		Traffic flow per road link, vehicle class and road type Fuel efficiency data	National vehicle ownership rates Population density Pre-calculated heat emissions for three vehicle classes	
Traffic diurnal variations	–	Separate profiles for six vehicle classes	One profile (all classes)	One profile (all classes)	

100 W m⁻² at 500 m resolution) towards the centre of London

2. Different spatial features between models, such as enhancements around major roads
3. The overall pattern of Q_F , which is more intense in the city centre

The Fractions Skill Score (FSS; Roberts and Lean 2008) is a metric that allows forecast skill at different spatial scales to be quantified in light of spatial errors and bias. Originally applied to compare forecast and observed precipitation fields (e.g. Roberts and Lean 2008; Mittermaier and Roberts 2010), it has also been used to evaluate modelled cloud brightness temperatures (e.g. Griffin et al. 2017) and volcanic ash plumes (Harvey and Dacre 2016). The FSS is calculated for different regimes of interest (e.g. Q_F intensity—see Table 2—rainfall rate or ash concentration) to highlight different aspects of model behaviour. In this study, regimes are selected based on Q_F magnitude and typically occupy certain regions of the city. For example, $Q_{F,B}$ emissions are consistently higher in the city centre; therefore, the upper regimes are found here (Fig. 3a). For a given regime, FSS is estimated as follows:

1. Grid cells belonging to the regime are tagged.
2. A neighbourhood of fixed size surrounding every grid cell in the model domain is chosen.
3. Spatial accuracy is assessed in each neighbourhood (Roberts, 2008). For the j^{th} neighbourhood:
 - (a) The fraction of tagged cells in the neighbourhood is calculated in the candidate model grid (M_j).
 - (b) The corresponding value is calculated in a reference grid (O_j), which may be another model or observations.
4. The fractions are calculated in all N neighbourhoods and combined into the FSS:

$$FSS = 1 - \frac{\sum_{j=1}^N (O_j - M_j)^2}{\sum_{j=1}^N O_j^2 + \sum_{j=1}^N M_j^2} \quad (1)$$

The FSS is calculated for multiple regimes to evaluate different aspects of the model output. Ordinarily, neighbourhoods of varying size are evaluated to determine spatial accuracy across different scales. Spatial resolution

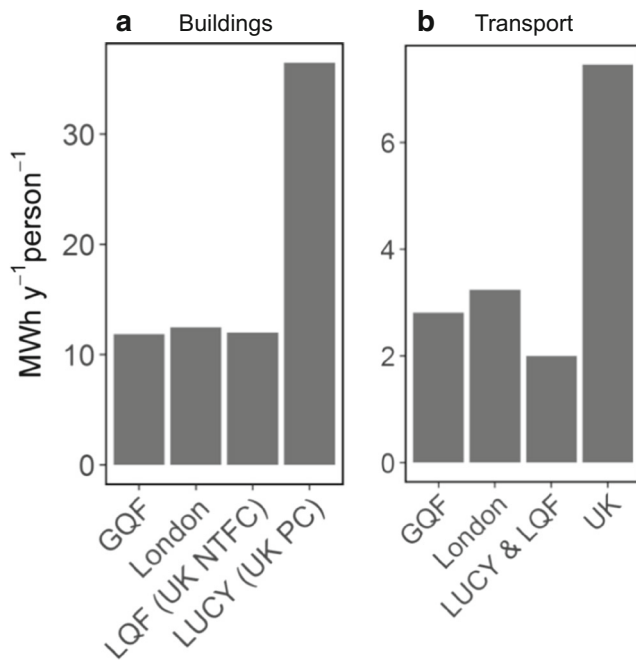


Fig. 2 Per-capita annual energy emissions for **a** buildings and **b** transport estimated by GQF, LQF, and LUCY. LUCY and LQF respectively use national primary consumption (PC) and non-transport final consumption (NTFC) (IEA, 2016) for building energy consumption and both estimate transport emissions based on vehicle ownership. London-specific energy consumption values (DECC, 2014) are shown for comparison

rather than scale is of interest in this work, so a single-pixel neighbourhood is used with model outputs with progressively coarser resolutions.

The FSS ranges from 0 to 1, where $FSS < 1$ indicates poorer skill but not the underlying reasons for it. The frequency bias (the between-models ratio of grid cell counts in a regime) is therefore calculated separately to aid in interpretation. A lower limit, FSS_{useful} , was defined with the FSS to signify whether the candidate model gives an informative (“useful”) prediction of the reference:

$$FSS_{useful} = 0.5 + f/2 \quad (2)$$

where f is the overall fraction of tagged cells in the reference grid, and 0.5 is the likeliest outcome if the forecast is random.

FSS_{useful} is half-way between the two.

The FSS provides an intuitive measure of overlap between models. The impact of spatial consistency is visualised using three examples (Fig. 4b–d) wherein the two models produce similar numbers of cells within a regime (low frequency bias) but with varying degrees of spatial consistency.

3 Results

LQF and GQF were run for Greater London at 500 m spatial resolution, consistent with finer-scaled global population datasets (e.g. GPWv4). Results are compared on a typical working day: Tuesday, 5 May 2015 for each component of Q_F for each hour of output. Model outputs are compared primarily at the 500-m base spatial resolution, with coarser resolutions considered by spatially averaging the outputs in post-processing. A maximum cell size of 5 km maintains a sample size over 50 grid cells in the 60×45 km domain.

Skill at different model resolutions from 500 m to 5 km is compared during night-time (00:00–01:00 UTC) and daytime (06:00–07:00 UTC for transport, 11:00–12:00 UTC otherwise) to capture low/domestically dominated and high/non-domestically dominated emissions, respectively.

3.1 Selection of regimes

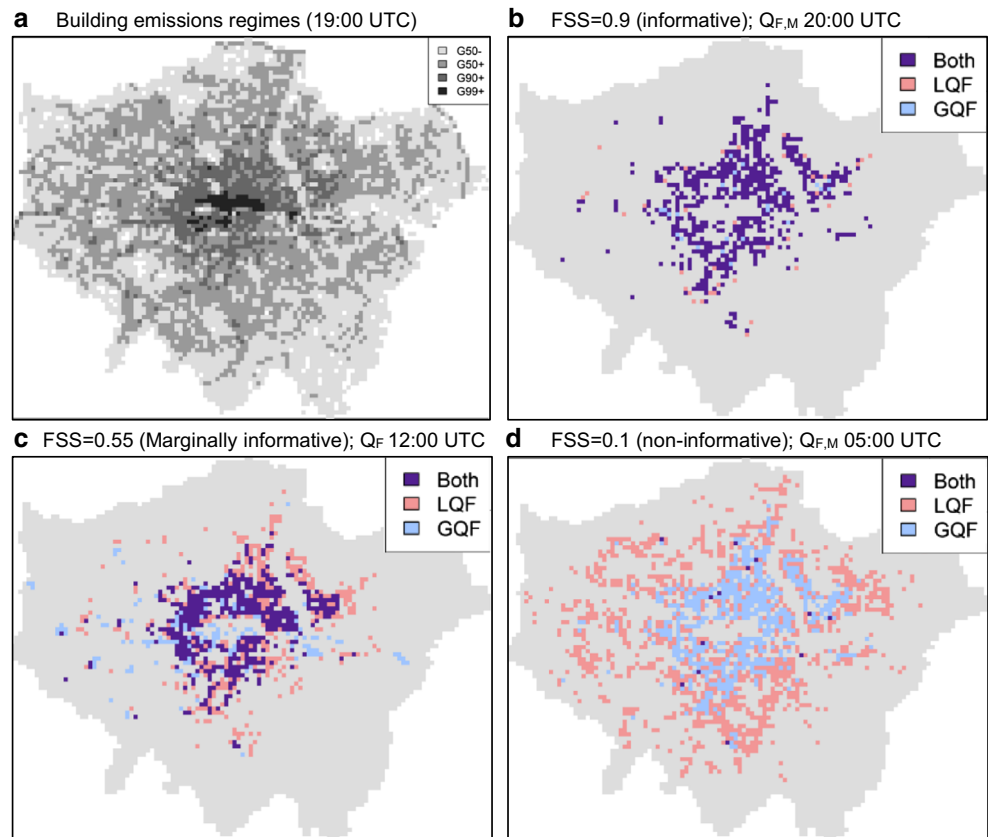
Q_F intensity regimes (defined in Table 2) are selected based on thresholds [$W m^{-2}$] applied to both model grids. Thresholds are estimated using quantiles of the emission intensity found in the GQF output grid. The resulting FSS therefore indicates if emissions are of the correct magnitude, and whether emissions of a given magnitude occur in the correct locations. Separate thresholds are calculated for each time step, Q_F component and spatial resolution.

The regimes are chosen such that the No- Q_F and G99+ cases, respectively, evaluate spatial intermittency and extremes, while the G50–, G50+ and G90+ regimes evaluate whether the bulk of Q_F values are predicted and assigned correctly, indicating if emissions are skewed higher or lower

Table 2 Definitions of anthropogenic heat flux (Q_F) intensity regimes used in model comparison, with corresponding Q_F values to illustrate the magnitude of each threshold in terms of GQF total Q_F and the proportion of the total GQF hourly energy emitted from each regime. Ranges are used because thresholds are re-evaluated at each time step

Regime	Description	Threshold for GQF total Q_F [$W m^{-2}$] (median)	% of GQF total hourly energy emitted
No- Q_F	GQF emissions $< 0.005 W m^{-2}$	–	–
G99+	Upper 1% of non-zero GQF values	26.7–90.6 (47.2)	6.3–11.5
G90+	Upper 10% of non-zero GQF values (excludes G99+)	8.4–25.0 (18.2)	23.4–26.1
G50+	Upper 50% of non-zero GQF values (excludes G90+)	3.3–9.4 (7.2)	46.6–52.2
G50–	Lower 50% of non-zero GQF values	–	15.8–18.2

Fig. 3 Examples of **a** where each emissions regime (Table 2) falls during a GQF time step (building emissions 5 May 2015 at 19:00 UTC), and maps that yield **b** high (informative), **c** medium (marginally above FSS_{useful}) and **d** low (non-informative) FSS values for the G90+ regime. Purple cells indicate where LQF and GQF emissions coincide, and pink and blue cells show mismatches between LQF and GQF (respectively) emissions. Note that different times and Q_F components are used to demonstrate each mismatch



than in GQF. The No- Q_F regime is evaluated only for $Q_{F,T}$ because the two models use different datasets to estimate spatial intermittency for transport, whereas population data (with zero values in identical locations) is used for other components. Skill in a given Q_F regime is considered to be informative if $FSS > FSS_{\text{useful}}$.

3.2 City-wide results

3.2.1 Day-to-day variations

GQF uses historical demand data to estimate daily building energy consumption, while LQF makes predictions based on ambient temperature (Sect. 2.1.2), and it is informative to establish the magnitude of the differences between these values (although beyond the scope of the work to resolve them). Based on model runs for all of 2015, the LQF/GQF city-wide daily building emissions ratio (Fig. 4e) varies from 0.6 to 1.5. The empirical relationship is bimodal about 1.1 and 0.8, and a seasonal breakdown (Fig. 5a–d) which arises from systematic under-estimates by LQF in the autumn and strong bimodality in winter compared to a central tendency about 1 in spring and summer. The mode at approximately 0.8 occurs because measured energy use generally exceeds that predicted by LQF when the temperature exceeds the LQF balance point

(Lindberg et al., 2013), which de-activates artificial heating in the model calculations.

These over/under-estimates are applied with equal weighting across all buildings in LQF, and therefore introduce frequency biases that reduce the FSS. The ratio is 0.97 on the study date (a typical mid-range value), so differences in building energy estimation methods are unlikely to confound the results of the spatial analysis.

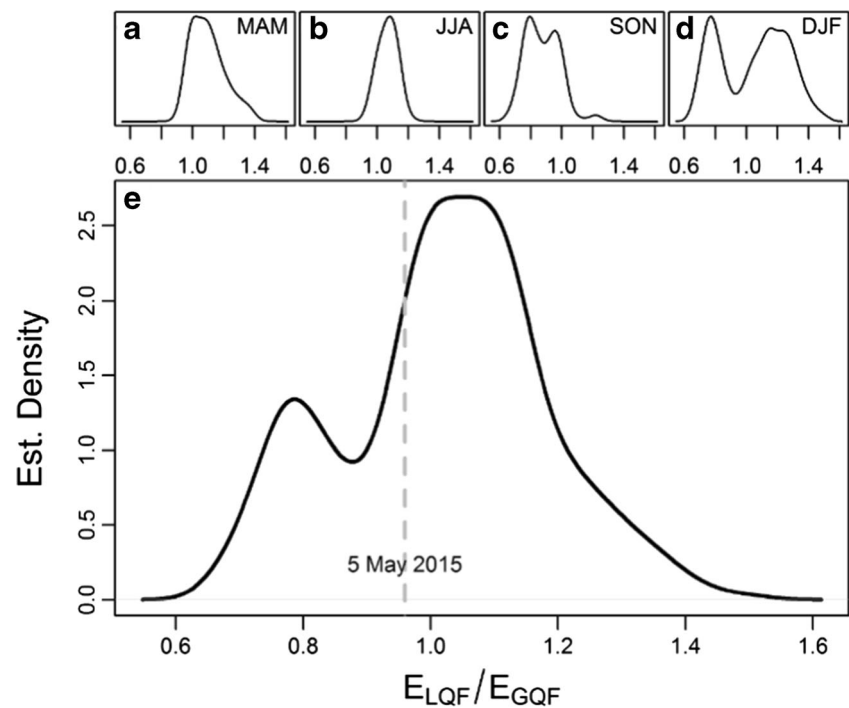
3.2.2 Diurnal variations

The variation in city-wide mean Q_F is consistent between models, with GQF and LQF (Fig. 3a, b) reaching maxima of 12 and 10 $W m^{-2}$ at 09:00 UTC (respectively) on 5 May 2015. Building energy dominates Q_F at all times, with transport emissions proportionally greatest between 06:00 and 07:00 UTC when building energy is still rising. Building emissions are greatest during the working day (non-domestic emissions) and evening (domestic emissions). The proportions shown in Fig. 5a vary spatially because GQF captures separate spatial patterns of transport, domestic and non-domestic emissions.

3.3 Total anthropogenic heat flux (Q_F)

The time-of-day variation of FSS for total Q_F (Fig. 6a) shows consistency between modelled emission peaks at night when

Fig. 4 Kernel density estimate of total daily building energy emissions in LQF (E_{LQF}) relative to GQF (E_{GQF}) over **a** spring, **b** summer, **c** autumn and **d** winter months of 2015, and **e** the whole year (same x-axis scale on each). Values of $E_{LQF}/E_{GQF} > 1$ indicate over-estimates by LQF with respect to GQF. Vertical dashed line represents the ratio found on 5 May 2015, the focus of this study ($n = 365$). LQF building emissions vary with air temperature while GQF emissions use empirical demand data

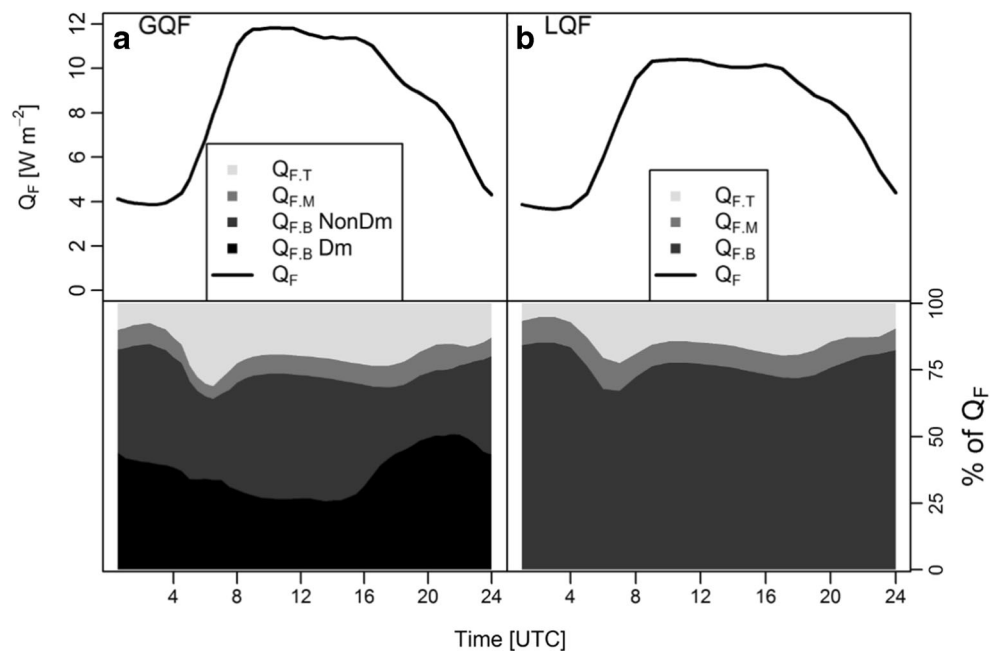


residential building emissions dominate Q_F . Skill falls during the day when transport and non-domestic emissions are greatest (Fig. 5). The G50– and G50+ regimes (50 and 40% of the area, respectively) are predicted with informative skill all day, but skill is lower during daytime; G90+ (9% of the area) is predicted informatively for much of the day but falls below the threshold of what is considered informative at 06:00, 07:00 and 15:00 UTC. The G99+ regime, which covers 1% of areas with non-zero Q_F , is predicted with negligible skill at all times.

Corresponding frequency biases at 01:00 and 07:00 (Table 3, rows 1 and 2) show that LQF predicts no emissions in the G99+ regime, hence the absence of skill. Biases in the G90+ regime are consistent at both times; hence, the reduction in FSS at 07:00 is likely caused by a change in the spatial Q_F distribution coinciding with increased traffic activity.

Progressively coarsening the spatial resolution, r , to 5 km (Fig. 6b) at 07:00 raises the skill of the G90+ regime to an informative level at 1 km. There is no such effect on the G99+ skill because the fixed boundaries of the model grid leads

Fig. 5 Time series showing the time evolution of (upper) mean total Q_F for all of London on 5 May 2015 and (lower) the proportion contributed by each Q_F component in **a** GQF and **b** LQF runs. GQF further separates building emissions into domestic (Dm) and non-domestic (NonDm) building emissions



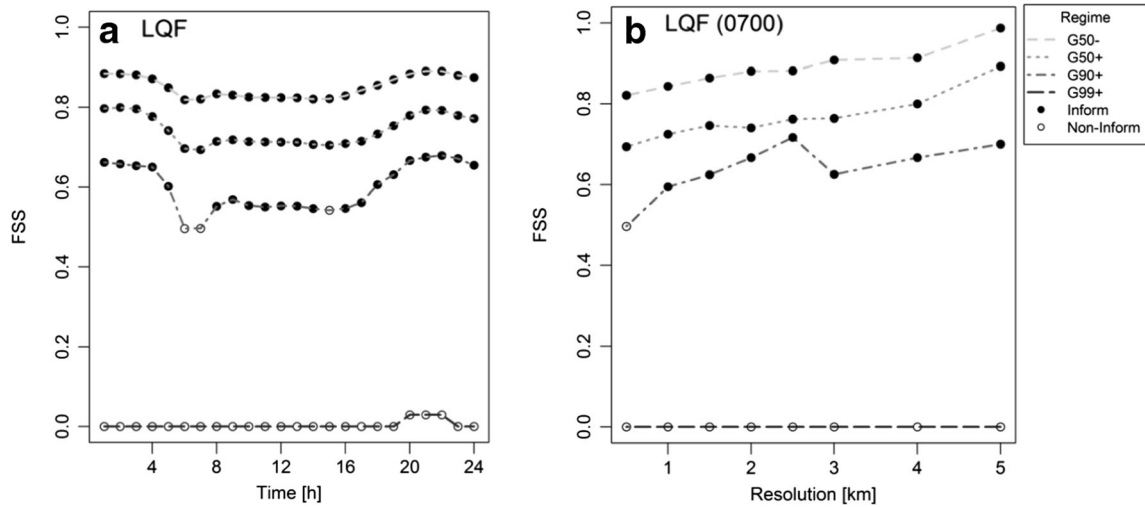


Fig. 6 Fraction skills scores (Eq. 1) for total Q_F in Greater London (500 m resolution) based on LQF and GQF-Standard by **a** time of day on 5 May 2015, and **b** spatial resolution at 07:00 UTC. One line shown

per Q_F intensity regime (Sect. 3.1). Darker lines indicate higher Q_F ; filled points denote informative FSS levels according to Eq. 2, and hollow points non-informative

higher emissions to consolidate to different regions of the domain as r reaches 5 km.

3.4 Transport emissions ($Q_{F,T}$)

The same approach is used to evaluate $Q_{F,T}$ emissions, with No- Q_F regime included here to indicate whether the spatially intermittent structure of the road network (Fig. 1f) is predicted adequately by LQF.

None of the $Q_{F,T}$ regimes are predicted informatively at any time of day (Fig. 7a). LQF greatly underestimates the frequency of grid cells in the No- Q_F , G90+ and G99+ regimes

(Table 3, rows 3 and 4). The frequency of emissions in the G50-regime is overestimated by 45%, though the G50+ regime frequency is within 15% of the correct value. This produces excessive $Q_{F,T}$ away from roads and underestimates emissions near major roads by spreading the available energy too thinly, reflecting how the road network structure differs from the residential population density (Fig. 1c). Coarsening the resolution to 5 km does not improve skill to an informative level except in the No- Q_F regime, which improves because emissions occur in all 5 km grid cells.

An improved spatial distribution was sought by normalising the available daily transportation energy to that

Table 3 Frequency biases (Candidate/Reference) in each emissions regime, by Q_F component (Sect. 1) and model variant (Sect. 2.1.1). Values greater than 1 indicate over-prediction by the candidate, and values under 1 indicate under-prediction. Non-informative entries are omitted. Table 2 defines the emissions regimes. Note that daylight savings time was in force on the date of the study

Candidate	Reference	Flux	Time (UTC)	Frequency bias by emissions regime				
				No- Q_F	G50-	G50+	G90+	G99+
LQF	GQF	Q_F	01:00	–	0.969	0.961	1.34	0
LQF	GQF	Q_F	07:00	–	0.985	0.956	1.29	0
LQF	GQF	$Q_{F,T}$	01:00	0.393	1.45	1.08	0.0486	0
LQF	GQF	$Q_{F,T}$	07:00	0.274	1.45	1.14	0.122	0
LQF-Paved	GQF	$Q_{F,T}$	01:00	0.151	1.01	1.76	0.237	0
LQF	GQF	$Q_{F,B}$	01:00	–	0.97	1.01	1.31	0
LQF	GQF	$Q_{F,B}$	07:00	–	0.971	1.01	1.31	0
GQF-Simplified	GQF	$Q_{F,B}$	01:00	–	0.975	0.994	1.24	0.429
GQF-Simplified	GQF	$Q_{F,B}$	07:00	–	0.975	0.994	1.23	0.422
LQF	GQF-Simplified	$Q_{F,B}$	01:00	–	0.996	1.03	1.08	0.0476
LQF	GQF-Simplified	$Q_{F,B}$	07:00	–	0.994	1.03	1.09	0.0469
LQF	GQF	$Q_{F,M}$	01:00	–	0.894	1.03	1.24	3.36
LQF	GQF	$Q_{F,M}$	07:00	–	0.683	0.802	2.25	14.6
LQF	GQF	$Q_{F,M}$	12:00	–	0.879	1.08	1.37	0

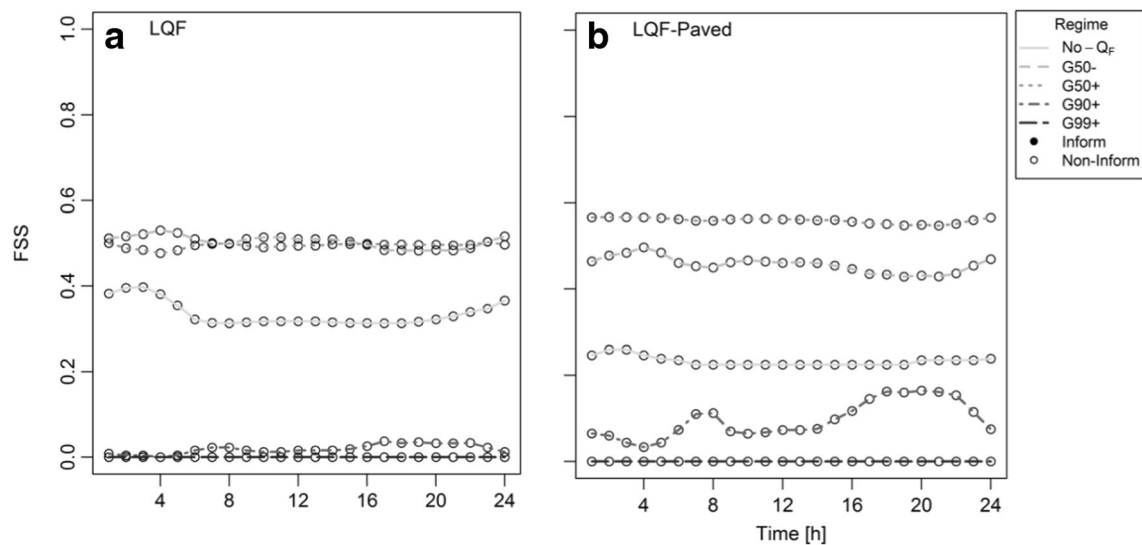


Fig. 7 As Fig. 6a, but for transport emissions using **a** LQF and **b** LQF-Paved, where daily emissions match GQF and emissions are assigned using paved area fraction rather than residential population

in GQF and using remote-sensed paved area fraction data in place of population density. This gives rise to the LQF-Paved configuration, which increased the FSS only marginally at 500 m resolution (Fig. 7b) because paved areas (Fig. 1g) resemble an amalgam of population and road network (Fig. 1c, e, f) rather than just the road network. As with the standard LQF configuration, the LQF-Paved frequency biases at 01:00 UTC (Table 3, row 5) show strong under-representation in the No- Q_F , G90+ and G99+ regimes, and over-prediction of the G50+ regime is strengthened. There is negligible bias in the G50- regime, so poor FSS here is caused by a lack of spatial consistency.

3.5 Building emissions ($Q_{F,B}$)

$Q_{F,B}$ is predicted informatively by LQF at all times of day in the G50-, G50+ and G90+ regimes (Fig. 8a). Frequency biases at 01:00 and 07:00 UTC are under 3% in the G50- and G50+ regimes, and the G90+ regime is overestimated by 31% (Table 3, rows 6 and 7). As with total Q_F , the $Q_{F,B}$ skill reduces during daytime when non-domestic emissions dominate in GQF, but not to the extent that skill falls below an informative level. Coarsening the spatial resolution to 5 km (not shown) does not increase the FSS in the G99+ regime.

The structure of the G99+ regime may arise from three aspects of the GQF input data that are not captured by LQF:

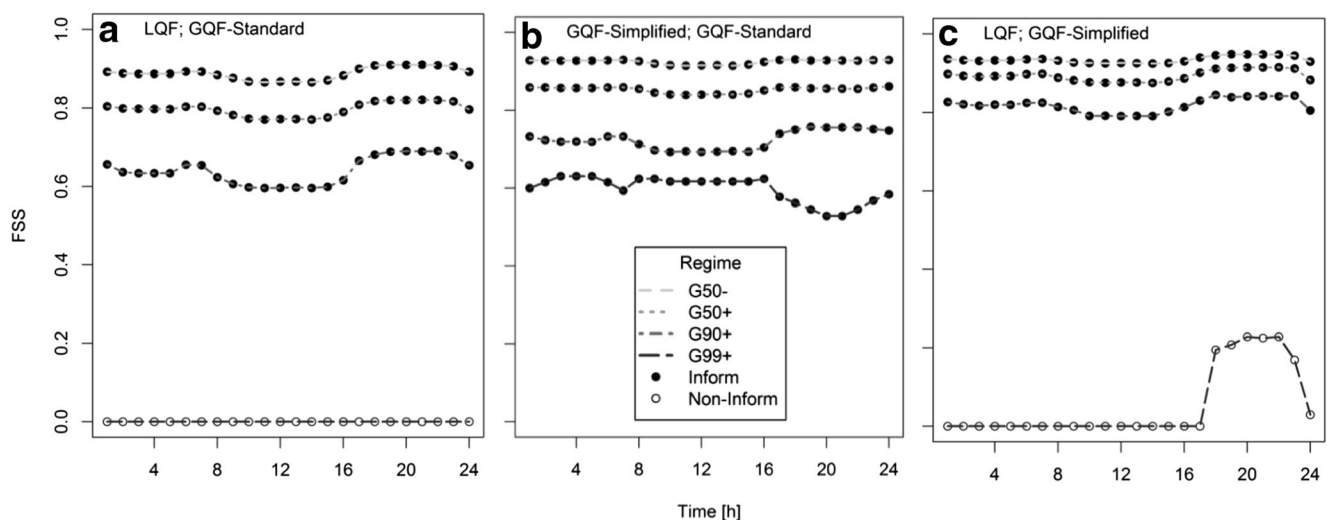


Fig. 8 As Fig. 6a, but showing time series of FSS for building emissions skill (500 m spatial resolution) for **a** LQF and **b** GQF-Simplified against GQF-Standard and **c** LQF against GQF-Simplified

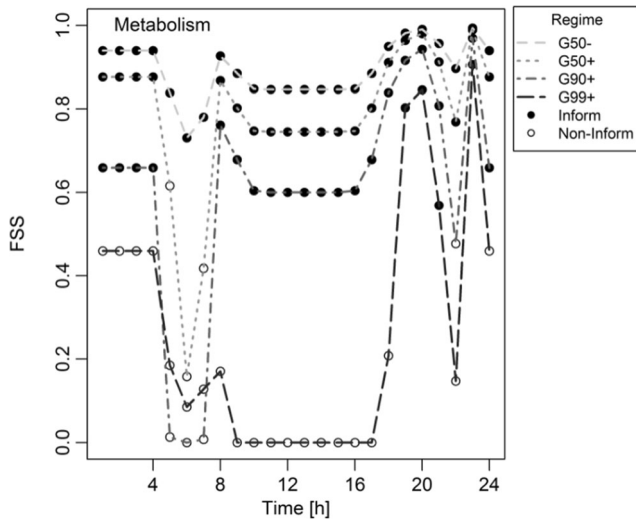


Fig. 9 As Fig. 6a, but showing FSS for metabolic emissions skill from LQF compared with GQF-Standard

1. Hour-to-hour energy consumption differs strongly between models because GQF uses sector-specific values and diurnal profiles.
2. The sector-specific building energy consumption datasets loaded into GQF are provided in spatially resolved form and generally show greater non-domestic consumption towards the city centre. In contrast, a city-wide consumption value is used in LQF.
3. Non-domestic energy consumption is disaggregated to the required spatial resolution using workday population, which has a different spatial structure to residential population (Fig. 1c, e). LQF disaggregates all energy consumption using residential population.

Hourly total building emissions in the two models were found to differ by less than 7%, suggesting (1) is not the primary cause of the lack of skill.

The GQF-Simplified configuration uses city-wide energy consumption totals instead of spatially resolved input

files to assess the effect of (2) on the G99+ regime. The resulting FSS (Fig. 8b) represents the skill of GQF-Simplified at predicting GQF. The FSS across regimes ranges from 0.5 (G99+) to 0.95 (G50-) and remains at informative levels all day, although it should be noted that the G99+ regime is only marginally above the FSS_{useful} threshold in the evening. The effect of GQF-Simplified is to reduce energy use in the city centre and the abundance of grid cells in the G99+ regime by 57–58% (Table 3, rows 8 and 9) at 01:00 and 07:00. This energy redistributes to grid cells in the G90+ regime, with a 23% increase in abundance. This corresponds to the energy from each lost G99+ cell being spread over 3.6 G90+ cells.

LQF was compared to GQF-Simplified to evaluate (3). The G99+ regime of GQF-Simplified is still predicted with negligible skill by LQF (Fig. 8c), albeit with minor improvements from 17:00 to 00:00 UTC, again caused by an absence of cells in the G99+ regime (Table 3, rows 10 and 11). This indicates the lack of skill in upper regimes is a combination of factors (2) and (3).

3.6 Metabolic emissions ($Q_{F,M}$)

GQF estimates night-time $Q_{F,M}$ using residential population and daytime emissions using workday population, with transitional periods during morning and evening and increased daytime metabolic activity per person. LQF uses residential population data at all times of day, and transitions to (from) increased metabolic activity during the morning (evening).

Workday and resident population differences lead to lower skill during the day than at night (Fig. 9), and the FSS rapidly worsens to non-informative levels at 05:00–07:00 and 22:00 UTC in the G50+, G90+ and G99+ regimes because day/night transitions take place differently in the two models. All regimes except the G99+ case are predicted with informative skill outside of transitional periods, and G99+ is predicted skilfully at 19:00–21:00 and 23:00.

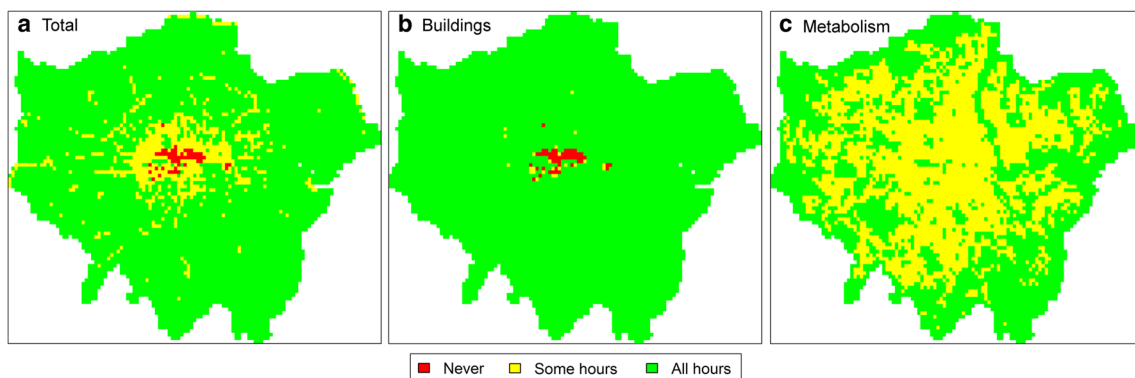


Fig. 10 Grid cell skill rating, indicating whether the regime(s) within which each pixel resided during 5 May 2015 were predicted with useful skill at all hours, some hours or never. Results shown for **a** Total Q_F , **b** $Q_{F,B}$ and **c** $Q_{F,M}$. Transport emissions are omitted because no useful skill was found (Fig. 7)

Non-informative skill in the $Q_{F,M}$ G99+ regime (Table 3, rows 12 to 14) is caused by frequency biases arising from different assumptions regarding resting metabolic rates and work schedules between models. The G99+ regime is over-predicted by 3.3 times at night because LQF assumes each person emits 75 W while GQF assumes 64.3 W. LQF predicts 14.6 times more cells at 07:00 UTC because its transition to active metabolic rates begins earlier than in GQF, which is conditioned on work/home rather than sleep/wake. At 12:00 UTC, transitions are complete and workday population dominates in GQF and LQF predicts zero G99+ cells, reflecting the high localised density of the workday population. Coarsening the spatial scale to 5 km (not shown) did not improve skill in the G99+ regime.

Night-time disagreement between models is trivial to resolve by adopting a consistent resting metabolic rate, but the differences during transitions between rest, wakefulness and work reflect the different levels of model detail.

3.7 Spatial variation of skill

Spatial variation in LQF skill is visualised by calculating the proportion of the day for which each grid cell resides in a regime with informative skill. This reliability (Fig. 10) is labelled as “consistently informative” ($FSS > FSS_{\text{useful}}$ during all hours), “intermittently informative” (some hours) or “poor” (never).

Total Q_F is intermittently informative in approximately 90% of the city area (Fig. 10a). Areas around major roads in the north, east and west of the city are reduced to being intermittently informative where the unskilled transport emissions are the strongest. The city centre contains the poorly predicted G99+ regime, which is dominated by building and transport emissions.

For $Q_{F,B}$ (Fig. 10b), over 98% of the city falls into regimes predicted consistently, and the G99+ regime at the centre is predicted without skill. The difference between Fig. 10a, b highlights the effect of roads on overall skill. Areas near the centre are intermittently informative because some grid cells fall into different regimes over the day. Metabolism (Fig. 10c from LQF is intermittently informative in the central half of the city where the workday population dominates and consistently informative in the residential outer 50%). $Q_{F,T}$ is not included because skill is consistently poor in all cases.

4 Discussion

4.1 Skill, emissions and energy

The total energy available for building emissions deviated between LQF and GQF on days with a mean temperature over 12 °C because LQF assumes no heating occurs in this regime.

The study date represented good agreement between models, and the bias introduced by this error reduces the FSS on other dates. It is stressed, however, that other prediction methods or empirical demand data could be used with the LQF approach instead.

FSS and area coverage are not related to the total energy contained within a regime (Table 2 contains a full breakdown of Q_F intensities and energy partitioning). The G50+ regime contains ~50% of energy, G90+ around 25% and G99+ approximately 10%. The areal extent of a regime therefore does not reflect its energetic significance, and obtaining good skill in the upper regimes may be more important for urban energy balance considerations if the focus is on high spatial resolution or hot spots.

4.2 Spatial accuracy

Total city-wide emissions in each component are consistent between models. LQF reproduced much of the GQF $Q_{F,B}$ spatial variability with an informative level of skill in most areas but was unable to accurately reproduce the city centre hot spots present in GQF output, with energy instead spread elsewhere. Accurate hot spot prediction requires workday population data, spatially constrained energy consumption data and the ability to discriminate between domestic and non-domestic emissions:

1. Spatially resolved domestic and non-domestic building energy consumption constrain emissions within different regions of the city.
2. Workday and residential populations indicate likely fine-grained building energy demand patterns during day and night.
3. Separate diurnal profiles for the energy consumption datasets correlate emissions in different areas to particular times of day.

LQF predicted spatial variations of $Q_{F,T}$ poorly in all regimes and times of day, smoothing out emissions over unsuitably large areas at resolutions as coarse as 5 km even when disaggregated using paved area fraction instead of population density. A road network map is required to address this, and the use of crowd-sourced vector data such as OpenStreetMap (2017) represents a potential avenue (subject to coverage) if assumptions are made about the division of traffic between major and minor roads.

$Q_{F,M}$ is predicted with informative levels of skill at night if accurate assumptions are made about per-person emissions. Differences in assumed resting and working times cause large transient losses in skill in the morning and evening. Daytime $Q_{F,M}$ predictions in LQF were predicted with poorer skill because LQF did not have access to workday population data, although only the hot spot regime fell below an informative level of skill during daytime.

Total Q_F skill reflects that of the individual components. LQF is non-informative at all times of day in the city centre because of building-related hot spots and non-informative during some hours of the day near major roads and dense parts of the road network. Since transport contributes less energy than buildings, consistently informative total Q_F skill can be obtained by coarsening the spatial resolution from 500 m to 1 km.

GQF is likely to more accurately represent true Q_F emissions than LQF; however, in this study, we cannot state how closely it matches reality.

5 Conclusions

A simple model (LQF) based principally on residential population and national statistics has limited accuracy at spatial resolutions from 0.5 to 5 km when compared with the output of a more detailed model (GQF), which uses city-specific parameters and distinguishes different energy uses:

- At the whole-city scale, building and road emissions were within $\pm 40\%$ of city-specific values, for individual days, with building emissions underestimated on warmer days.
- Elevated inner-city emissions, dominated by buildings, were displaced by LQF and hot spots were missed entirely. This is attributed to workday population and non-domestic energy use patterns.
- Outer-city emissions were replicated reliably by LQF as they are dominated by domestic buildings.
- Transport emissions were predicted poorly throughout the city because population and paved area fraction data attributed emissions across too great an area.
- Metabolic emissions were captured skilfully by LQF except during transitions between rest and activity.

We recommend that if detailed modelling is impractical because of limited input data, simple models based on residential population density patterns must be used conservatively:

- Resolutions no finer than ~ 1 km should be used to mitigate the effects of a lack of population movement (e.g. from home to work) being modelled.
- Transport emissions should be based on road network maps rather than a proxy, especially where major orbital and trunk roads displace traffic volume from dense populations.
- The relation between temperature and energy use should be evaluated for each study city. Amongst other improvements to LUCY (now LQF), through this work, optional extra parameters have been added to the LQF software to permit this.

- Errors arising from misplaced emission hot spots should still be expected despite these measures.

As new techniques are developed to obtain Q_F (e.g. Chrysoulakis et al. 2016), and high-resolution urban modelling uses Q_F as an input (e.g. Loridan et al. 2010; Chen et al. 2011; Bohnenstengel et al. 2014; Best and Grimmond 2016), estimation methods must be evaluated objectively so that their appropriateness can be judged and their limitations addressed. Given the challenge of obtaining all the data necessary to run a detailed model like GQF, enriching a simpler LQF-type model with complementary data may be a more fruitful way of improving the quality of predictions.

Despite extensive inputs, the models discussed here are essentially static and do not explicitly consider the effect of localised or widespread disruptions to human activity. Developing methods to emulate the dynamics of human behaviour therefore is essential so that spatially heterogeneous Q_F predictions can be made (Barlow et al. 2017). In turn, this will support urban energy balance and surface-atmosphere interaction modelling at progressively higher spatial and temporal resolutions.

Acknowledgements The authors would like to extend thanks to Fredrik Lindberg for advice and the City of London Barbican Estate for providing the measurement site.

Funding information Financial support by the European Union H2020-EO-1-2014 Project 637519: URBANFLUXES, and the UK-China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. The UMEP toolkit is available for download at <http://www.urban-climate.net/umep>.

Appendix 1. Model parameters

The following provide the parameters and units used in the models (Tables 4 and 6) and data sources (Tables 5 and 7) used

Table 4 Parameters used in LQF model

LQF Parameter	Value	Unit
Vehicle speed	48,000	m h^{-1}
Heat emissions factors by vehicle type	Cars 25.92 Motorcycles 13.16 Freight 108.42	W m^{-1}
Balance point temperature	12	$^{\circ}\text{C}$
Building energy weighting at balance point	0.7	N/A
Proportion of vehicles in use	0.8	N/A
Individual metabolic emission	Sleeping 75 Working 175	W

Based on Allen et al. (2011) and Lindberg et al. (2013). Heating/cooling response curve parameters (not shown) for the UK are as in Lindberg et al. (2013)

Table 5 Data sources used to populate LQF database, any/all of which can be updated by the user of the software

LQF input data	Reference(s)	Year updated
National vehicle ownership by class	Cars, motorcycles and freight: Worldmapper (2006)	2002
Vehicle emission factors by class	Smith et al. (2009)	2001
National population	Central Intelligence Agency World Fact Book (2016)	2016
Final national energy consumption	IEA (2016)	2014
UN national income grouping	World Bank (see Lindberg et al. 2013)	2012
Metabolic heat and activity timings	Smith et al. (2009); Lindberg et al. (2013)	N/A

The newest available values (as of July 2017) were used for each data source

Table 6 Parameters used in GQF calculations

GQF Parameter	Value	Unit
Electric water heater efficiency	0.98	N/A
Gas water heater efficiency	0.85	N/A
Latent fraction of metabolic heat	0.3	N/A
Sensible fraction of metabolic heat	0.7	N/A
Vehicle age (relative)	5	Year
Fraction of domestic electricity used for water heating	0.139	N/A
Fraction of domestic gas used for water heating	0.27	N/A
Fraction of non-domestic electricity used for water heating	0.036	N/A
Fraction of non-domestic gas used for water heating	0.146	N/A
Natural gas heat of combustion	Net of latent heat 35.5 Gross 39.4	MJ kg ⁻¹
Petrol heat of combustion	Net 44.7 Gross 47.1	MJ kg ⁻¹
Diesel heat of combustion	Net 43.3 Gross 45.5	MJ kg ⁻¹

All taken from Iamarino et al. (2012) except for vehicle age, which was added in the current work and is assumed to be 5 years. Economy 7 electricity and crude oil parameters are omitted from this study

Table 7 Summary of data sources used in GQF, along with spatial and temporal resolution

GQF data source description	Spatial resolution	Temporal resolution	Reference
Domestic electricity use maps	Domestic: Lower-level super output area (LSOA)	Annual	DBEIS (2015b)
(Non-)domestic gas use maps	LSOA, mid-level SOA	Annual	DBEIS (2015a)
Non-domestic electricity use map	Borough level	Annual	London Datastore (2016)
Daily UK gas demand	N/A	Daily	National Grid (2016a)
Half-hourly UK electricity demand	N/A	30 min	National Grid (2016b)
Building energy load profiles: One diurnal cycle representing all weekdays, one each for Saturday, Sunday	N/A	10 year rolling-average	Exelon UK (2012)
Residential population (2011)	Census output area (OA)	Annual	ONS (2013)
Workday population (2011)	OA	10 year updates, estimates in between	ONS (2014)
Road link map w/ traffic count by vehicle type	Greater London	Infrequent updates	London Datastore (2014)
Traffic diurnal profiles by vehicle type	Greater London	Infrequent updates	Iamarino et al. (2012) model input data
Fuel consumption by different euro class, vehicle and road type	N/A	Unknown	DEFRA (1999)
Estimated calorific values of fuels	N/A	Unknown	DBEIS (2014)

Table 8 List of vehicle classes modelled in GQF, along with how the model gets information about the relative balance of petrol and diesel vehicles in each class

Vehicle type	Source of information on fuel mix	Assumed fuel mix
Motorcycle	Assumption	All petrol
Car	Input data (if available)	84% petrol and 16% diesel
Light goods vehicle	Input data (if available)	10% petrol and 90% diesel
Bus	Assumption	All diesel
Taxi	Assumption	All diesel
Rigid heavy goods vehicles	Assumption	All diesel
Articulated heavy goods vehicles	Assumption	All diesel

Assumed fuel mix values are taken from Iamarino et al. (2012). Assumed fuel mix is subject to change with regulatory and technological advances

as inputs to GQF and LQF. The data used to populate the LQF database of national statistics is listed in Table 5.

Appendix 2. Algorithm used to produce gridded population and transport energy data

(Dis)aggregation of population/energy consumption data

The disaggregation algorithm redistributes V_S , a set of scalar quantities attributed to a set S of “source” spatial units to a set T of target units that intersect them. These units are of arbitrary shape, and S and T may have different spatial extents to one another. Each target unit receives a share of the quantities V_S according to the area of source unit(s) intersected and the value of a numerical weighting specified for each target unit.

The process and rules to perform the disaggregation are outlined below. Each source unit is processed in turn, and its contributions to each target area are summed afterwards.

1. The n target units spatially intersecting the j^{th} source unit are identified. The i^{th} target unit has an overall area $A_{T,i}$ and intersects an area $a_{T,i}$. The n target units collectively overlap a fraction F of the overall source area $A_{S,j}$
2. The i^{th} target unit receives $V_{T,i}$. This is a fraction of $V_{S,j}$ determined by a weighting w_T assigned to each of the n target areas. Weightings may be provided externally, or the overall target unit area is used otherwise (i.e. $w_{T,i} = A_{T,i}$):

$$V_{T,i} = \frac{w_{T,i}}{W_{S,j}} V_{S,j} \quad (3)$$

The term $W_{S,j}$ represents the total weighting over the source unit. This is either:

- (a) Specified externally before processing, which allows $W_{S,j} \neq \sum_i w_{T,i}$, whereby $V_{S,j}$ is scaled up or down before being disaggregated.

- (b) Calculated from the sum of the individual weights, taking into account the fraction F to scale down $V_{S,j}$ if the target units do not completely cover the source unit:

$$W_{S,j} = \frac{1}{F} \sum_i w_{T,i} \quad (4)$$

- (c) If $W_{S,j} = 0$, it is overridden with $W_{S,j} = \frac{1}{n}$.

3. A target unit may intersect m source units. In this event, the i^{th} target area weighting w_i used in Eq. 3 is first scaled by the proportion of the target unit area intersected by the source unit:

$$w'_{T,i} = \frac{a_{T,i}}{A_{T,i}} w_{T,i} \quad (5)$$

The process therefore performs aggregation if a target unit bounds multiple source units and disaggregation if the source areas are larger than the target unit. The resulting value of $V_{S,j}$ is conserved across the n target units (i.e. $\sum_i V_{T,i} = V_{S,j}$)

Aggregation of road network data (GQF)

The London Atmospheric Emissions Inventory (LAEI; London London Datastore, 2014) supplies a detailed road segment map, with annual average daily traffic (AADT) specified on a per-segment and per-vehicle class (Table 8) basis. These segments are vector lines and are infinitesimally thin, so the data are transformed to fuel consumption in each output polygon:

Total fuel consumption F in an output polygon of area A is obtained by summing contributions from each of the m road segments passing through it. The total fuel consumption f_v for vehicle of class v is calculated as in Eq. 6, where l_i is road segment length in kilometres, $D_{i,v}$ is the annual average daily traffic (AADT) of the vehicle class on the road segment and ε_v is the fuel consumption of the

vehicle class (kg km^{-1} ; based on EURO-II derived estimates on urban roads (DEFRA, 1999).

$$f_v = \sum_i^m \varepsilon_v l_i D_{i,v} \quad (6)$$

$$Q_{F,T} = \frac{10^6}{48 \cdot A} \sum_j^n f_v q_v \quad (7)$$

Mean half-hourly $Q_{F,T}$ is estimated (Eq. 7) by summing across all n vehicle types and applying the heat of combustion. The available data splits vehicles into petrol and diesel variants so the appropriate heat of combustion q_v from Iamarino (2012) is applied (44.7 and 47.1 MJ kg^{-1} for petrol, net and gross of water vapour and 43.3 and 45.5 MJ kg^{-1} for diesel, net and gross of water vapour). Where data regarding the fuel split is not available, assumptions are made (Table 8). Half-hourly variations are incorporated by applying a diurnal scaling factor, which also varies from day to day to reflect reduced weekend traffic flow.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allen L, Lindberg F, Grimmond CSB (2011) Global to city scale urban anthropogenic heat flux: model and variability. *Int J Climatol* 31(13):1990–2005. <https://doi.org/10.1002/joc.2210>
- Barlow J, Best M, Bohnenstengel SI, Clark P, Grimmond S, Lean H, Christen A, Emeis S, Haeffelin M, Harman IN, Lemonsu A, Martilli A, Pardyjak E, Rotach MW, Ballard S, Boutle I, Brown A, Cai X, Carpentieri M, Coccal O, Crawford B, Di Sabatino S, Dou J, Drew DR, Edwards JM, Fallmann J, Fortuniak K, Gornall J, Gronemeier T, Halios CH, Hertwig D, Hirano K, Holslag AAM, Luo Z, Mills G, Nakayoshi M, Pain K, Schlünzen KH, Smith S, Souhac L, Steeneveld G-J, Sun T, Theeuwes NE, Thomson D, Voogt JA, Ward HC, Xie Z-T, Zhong J (2017) Developing a research strategy to better understand, observe and simulate urban atmospheric processes at kilometre to sub-kilometre scales. *Bull Am Meteorol Soc* 98(10):ES261–ES264. <https://doi.org/10.1175/BAMS-D-17-0106.1>
- Best MJ, Grimmond CSB (2016) Investigation of the impact of anthropogenic heat flux within an urban land surface model and PILPS-urban. *Theor Appl Climatol* 126(1):51–60. <https://doi.org/10.1007/s00704-015-1554-3>
- Bohnenstengel SI, Hamilton I, Davies M, Belcher SE (2014) Impact of anthropogenic heat emissions on London's temperatures. *Q J R Meteorol Soc* 140(679):687–698. <https://doi.org/10.1002/qj.2144>
- Bueno B, Pigeon G, Norford LK, Zibouche K, Marchadier C (2012) Development and evaluation of a building energy model integrated in the TEB scheme. *Geosci Model Dev* 5(2):433–448. <https://doi.org/10.5194/gmd-5-433-2012>
- Center for International Earth Science Information Network, C. C. U (2016) Gridded population of the world, version 4 (GPWv4): population density. Palisades, NASA Socioeconomic Data and Applications Center (SEDAC)
- Chen F, Kusaka H, Bornstein R, Ching J, Grimmond CSB, Grossman-Clarke S, Loridan T, Manning KW, Martilli A, Miao S, Sailor D, Salamanca FP, Taha H, Tewari M, Wang X, Wyszogrodzki AA, Zhang C (2011) The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *Int J Climatol* 31(2):273–288. <https://doi.org/10.1002/joc.2158>
- Chrysoulakis N, Heldens W, Gastellu-Etchegorry J-P, Grimmond S, Feigenwinter C, Lindberg F, Del Frate F, Klostermann J, Mitraha Z, Esch T (2016) A novel approach for anthropogenic heat flux estimation from space Geoscience and remote sensing symposium (IGARSS), 2016 I.E. International, IEEE
- CIA (2016) World Fact Book: National population, https://www.cia.gov/library/publications/the-world-factbook/rankorder/rawdata_2119.txt Accessed 31/10/2016
- DBEIS (2014) DUKES A.1-A.3: Estimated average calorific values of fuels. <https://www.gov.uk/government/statistics/dukes-calorific-values>. Accessed 15 July 2016
- DBEIS (2015a) Spatially resolved domestic and non-domestic gas consumption. <https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-gas-consumption>. Accessed 20 July 2016
- DBEIS (2015b) Spatially resolved domestic electricity consumption. <https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-electricity-consumption>. Accessed 20 July 2016
- DBEIS (2016) Sub-national electricity and gas consumption statistics, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/579203/Sub-national_electricity_and_gas_consumption_summary_report_2016.pdf accessed 01/02/2017
- DECC (2014) Total energy consumption 2005–2012, <https://www.gov.uk/government/statistical-data-sets/total-final-energy-consumption-at-regional-and-local-authority-level-2005-to-2010> accessed 24/05/2017
- DEFRA (1999) National atmospheric emissions inventory: UK emissions of air pollutants 1970 to 1999 <https://uk-air.defra.gov.uk/assets/documents/reports/empire/naei/annreport/annrep99/index.htm> accessed 10/10/2016
- DEFRA (2014) National Atmospheric Emissions Inventory: base 2013 fleet composition data, http://naei.defra.gov.uk/resources/rtp_fleet_projection_Base2013_v3.0_final.xlsx accessed 01/08/2016
- Dong Y, Varquez ACG, Kanda M (2017) Global anthropogenic heat flux database with high spatial resolution. *Atmos Environ* 150:276–294. <https://doi.org/10.1016/j.atmosenv.2016.11.040>
- Exelon UK (2012) Average profiling data per profile class (regression data evaluated at 10-year average temperatures). https://www.exelon.co.uk/wp-content/uploads/2012/01/Average_Profiling_data_201314_evaluated@10yearNET_v1.0.xlsx. Accessed 10 Aug 2016
- Flanner MG (2009) Integrating anthropogenic heat flux with global climate models. *Geophys Res Lett* 36(2):L02081. <https://doi.org/10.1029/2008GL036465>
- Griffin SM, Otkin JA, Rozoff CM, Sieglaff JM, Counce LM, Alexander CR (2017) Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? *Weather Forecast* 32(1):5–25. <https://doi.org/10.1175/waf-d-16-0098.1>
- Hamilton IG, Davies M, Steadman P, Stone A, Ridley I, Evans S (2009) The significance of the anthropogenic heat emissions of London's buildings: a comparison against captured shortwave solar radiation. *Build Environ* 44(4):807–817. <https://doi.org/10.1016/j.buildenv.2008.05.024>
- Harvey NJ, Dacre HF (2016) Spatial evaluation of volcanic ash forecasts using satellite observations. *Atmos Chem Phys* 16(2):861–872. <https://doi.org/10.5194/acp-16-861-2016>

- Iamarino M, Beevers S, Grimmond CSB (2012) High-resolution (space, time) anthropogenic heat emissions: London 1970–2025. *Int J Climatol* 32(11):1754–1767. <https://doi.org/10.1002/joc.2390>
- IEA (2016) Headline energy data. http://www.iea.org/media/statistics/IEA_HeadlineEnergyData_2016.xlsx. Accessed 23 May 2017
- Kikegawa Y, Genchi Y, Yoshikado H, Kondo H (2003) Development of a numerical simulation system toward comprehensive assessments of urban warming countermeasures including their impacts upon the urban buildings' energy-demands. *Appl Energy* 76(4):449–466. [https://doi.org/10.1016/S0306-2619\(03\)00009-6](https://doi.org/10.1016/S0306-2619(03)00009-6).
- Lindberg F, Grimmond CSB, Yogeswaran N, Kotthaus S, Allen L (2013) Impact of city changes and weather on anthropogenic heat flux in Europe 1995–2015. *Urban Climate* 4:1–15. <https://doi.org/10.1016/j.uclim.2013.03.002>.
- Lindberg F, Grimmond CSB, Gabey A, Huang B, Kent CW, Sun T, Theeuwes NE, Järvi L, Ward H, Capel-Timms I, Chang YY, Jonsson P, Krave N, Liu DW, Meyer D, Olofson KFG, Tan JG, Wästberg D, Xue L, Zhang Z (2018) Urban multiscale environmental predictor (UMEP)—an integrated tool for city-based climate services. *Environ Model Softw* 99:70–87. <https://doi.org/10.1016/j.envsoft.2017.09.020>
- London Datastore (2014). London Atmospheric Emissions Inventory (LAEI) 2013 - Supporting information: key GIS geographies and road traffic flows and vehicle-kilometres, <https://files.datapress.com/london/dataset/london-atmospheric-emissions-inventory-2013/2016-03-15T13:38:41/1.%20Supporting%20Information.zip> Accessed 07/03/2016
- London Datastore (2016) Borough-level electricity consumption (Database ID 31bcc41a-9d06-412a-9fd5-9de591e5e900). <http://data.london.gov.uk/dataset/electricity-consumption-borough>. Accessed 20 July 2016
- Loridan T, Grimmond CSB, Offerle BD, Young DT, Smith TEL, Järvi L, Lindberg F (2010) Local-scale urban meteorological parameterization scheme (LUMPS): longwave radiation parameterization and seasonality-related developments. *J Appl Meteorol Climatol* 50(1): 185–202. <https://doi.org/10.1175/2010JAMC2474.1>.
- Marconcini M, Heldens W, Del Frate F, Latini D, Mittra Z, Lindberg F (2017) EO-based products in support of urban heat fluxes estimation. Urban remote sensing event (JURSE), 2017 Joint, IEEE
- Mittermaier M, Roberts N (2010) Intercomparison of spatial forecast verification methods: identifying skillful spatial scales using the fractions skill score. *Weather Forecast* 25(1):343–354. <https://doi.org/10.1175/2009waf2222260.1>
- National Grid (2016a) Daily UK gas demand data. <http://marketinformation.natgrid.co.uk/gas/DataItemExplorer.aspx>. Accessed 10 Aug 2016
- National Grid (2016b) Half-hourly UK electricity demand. <http://www2.nationalgrid.com/UK/Industry-information/Electricity-transmission-operational-data/Data-Explorer>. Accessed 10 Aug 2016
- ONS (2013) England and Wales 2011 census data: usual resident population, <https://www.nomisweb.co.uk/census/2011/ks101ew> Accessed 1/3/2017
- ONS (2014) England and Wales 2011 census data: Workday population. <https://www.nomisweb.co.uk/census/2011/wd607ew>. Accessed 3 Jan 2017
- ONS (2016) 2011 Census aggregate data. Office for National Statistics UK Data Service (Edition: June 2016), doi: <https://doi.org/10.5257/census/aggregate-2011-1>
- OpenStreetMap contributors (2017) OpenStreetMap. from <https://www.openstreetmap.org>
- Pesaresi SFM (2015) GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015). J. R. C. J. European Commission
- Quantum GIS Development Team (2016) Quantum GIS geographic information system, Open Source Geospatial Foundation Project
- Roberts N (2008) Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol Appl* 15(1): 163–169. <https://doi.org/10.1002/met.57>
- Roberts NM, Lean HW (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon Weather Rev* 136(1):78–97. <https://doi.org/10.1175/2007mwr2123.1>
- Sailor DJ, Georgescu M, Milne JM, Hart MA (2015) Development of a national anthropogenic heating database with an extrapolation for international cities. *Atmos Environ* 118:7–18. <https://doi.org/10.1016/j.atmosenv.2015.07.016>.
- Sailor DJ, Lu L (2004) A top-down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas. *Atmos Environ* 38(17):2737–2748. <https://doi.org/10.1016/j.atmosenv.2004.01.034>.
- Smith C, Lindley S, Levermore G (2009) Estimating spatial and temporal patterns of urban anthropogenic heat fluxes for UK cities: the case of Manchester. *Theor Appl Climatol* 98(1):19–35. <https://doi.org/10.1007/s00704-008-0086-5>
- Worldmapper (2006) Worldmapper: passenger cars, freight and motorcycles <http://www.worldmapper.org/display.php?selected=31>. Accessed 31 October 2016